Short-Term Forecasting for Distributed Energy Resources Visibility in Distribution System Operations

Technical Brief — *Distribution Operations and Planning*

Abstract

Distribution utilities across the United States have varying levels of visibility into Distributed Energy Resources (DERs). As the penetration of solar photovoltaic (PV) systems and other DERs increase, DER visibility is becoming more critical to ensure safe and reliable operations of the distribution grid. Solutions are needed that will increase hosting capacity capabilities, enable improved operational decisions, inform future DER dispatch needs, and maintain distribution grid reliability. Short-term forecasting is a potential solution to improving DER visibility that is both cost-effective and scalable to adapt to the changing DER landscape. This document summarizes learnings from a study with a large US utility, where an advanced DER forecasting solution was evaluated against approximately two years of 15-minute resolution measurements from over 70 distributed PV installations within the utility's service territory.

Keywords

Short-term forecasting
DER integration
Distribution operations
Grid operations

Executive Summary

PRIMARY AUDIENCE: Distribution operators

SECONDARY AUDIENCE: Distribution planners, DER Aggregators

KEY RESEARCH QUESTION

The goal of this document is to give insight into the use of short-term forecasting to provide visibility of DER production, with an emphasis on practical solutions that can be adopted today by Distribution operators. This includes benchmarking of forecast accuracy across timescales relevant to operations (minutes to hours to days ahead) compared to direct measurement.

RESEARCH OVERVIEW

The analysis evaluated the use of advanced forecasting techniques to provide customer-level and aggregated estimates of distributed PV production without access to direct measurements. Forecast performance was evaluated against real-world distributed PV systems on timescales relevant to Distribution operations, ranging from minutes to days ahead.

KEY FINDINGS

- Short-term forecasting can provide DER visibility across timescales of interest to Distribution operations (minutes to days ahead) and without requiring DER measurement data.
- Forecast performance depends on accurate knowledge of DER system details, e.g., installed capacity [kW], PV panel orientation and tilt, and shading of PV panels from trees.
- Forecasts can provide DER visibility down to the customer-level, but forecasts are more accurate when aggregated to coarser granularities, e.g., feeder-level.
- Forecast accuracy can be improved if (historical) measurement data is available, even if measurements are only available at a subset of sites within a region.

WHY THIS MATTERS

The increasing penetration of DERs provide challenges to maintaining safe and reliable operations of the distribution grid. These challenges are exacerbated by limitations on visibility of DER production, which may in turn mask the true demand. This report provides insights into the use of short-term forecasting as a scalable and cost-effective solution to the lack of DER visibility for Distribution operators. The learnings here also have the potential to be aggregated to the transmission level, though this is not the focus of the paper.

HOW TO APPLY RESULTS

Distribution operators may use this report as a primer on short-term forecasting and its use in providing DER visibility. The results may also be used as benchmarks regarding the accuracy of distributed PV production forecasting as a function of forecast horizon, schedule, and granularity, e.g., site-level vs feeder-level forecasts.

EPRI CONTACTS: Mobolaji Bello, Technical Leader, mbello@epri.com
David Larson, Engineer/Scientist III, dlarson@epri.com

PROGRAM: Distribution Operations and Planning

Table of Contents

This technical brief was prepared by EPRI.

Introduction

Utilities across the United States have varying levels of visibility into Distributed Energy Resources (DERs). In some regions, utilities require monitoring equipment on mid- and large-sized installations, but not for smaller systems (e.g., <50 kW). Additionally, some existing monitoring capabilities do not distinguish between customers with and without DER and lack the ability to quantify the amount of customer DER production at any given time. The result is that Distribution operators have limited to no visibility into DER production and therefore limited knowledge of gross (aka 'native') load on a feeder (see Figure 1). As PV penetration increases, this lack of visibility is becoming more critical to ensure safe and reliable operation of the distribution system. By lacking observability in distribution system operations, utilities are negatively impacted by failed automation restoration schemes and limitations in the ability to utilize volt-var optimization (VVO). Solutions are needed that will increase host capacity capabilities, enable improved operational decisions, inform DER dispatch needs, and maintain distribution grid reliability. At the same time, these solutions must be cost-effective and scalable to adapt to the changing DER landscape.

The Challenge of DER Visibility

There are two main situations that result in lack of DER visibility for distribution system operations. First is lack of sufficient telemetry, either due to lack of monitoring equipment or the inability of existing equipment to directly measure DER production (e.g., meters that measure net load to/from the grid rather than gross load and DER production separately). Second is the situation in which there is sufficient telemetry, but there is too much lag–also referred to as latency–between the time of measurement and the time at which the data reaches operators (e.g., measurements recorded every 15-min, but only transmitted once per day) or the sampling frequency is not fast enough for the given application (e.g., one measurement per hour when 15-min data is required).

Forecasting as a Solution

Short-term forecasting offers a potential solution to the above challenges that is both cost-effective and scalable. The basic premise is to develop models to predict the current and future production of DERs without having access to real-time measurements from the DERs at time of prediction. These forecasts may serve as direct replacements for actual measurements in operational decisions and processes or be combined and transformed with other sources of information. For example, a forecast of the current DER production (i.e., a nowcast) could be combined with net meter data to estimate the "true" demand.

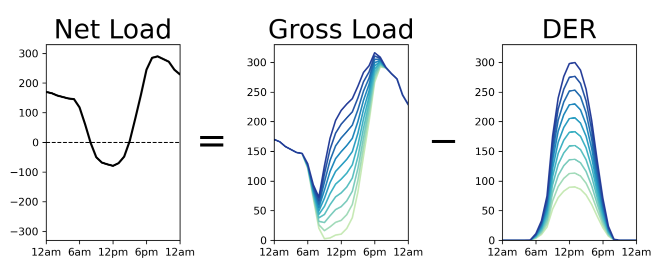

Figure 1. Illustration of how a net load profile can be the result of any number of gross load and DER combinations.

Forecasting Basics

This section briefly introduces key concepts used when discussing forecasts throughout the rest of the document.

Types of Forecasts

Short-term forecasts can be categorized into three general types:

- **Deterministic**: predict a numeric value, e.g., the total load will be 8 MW at 9am
- **Probabilistic**: predict a numeric value and its associated likelihood of occurrence, e.g., there is a 92% probability that the PV plant will produce at least 3 MW at 1pm
- **Event**: predict whether an outcome occurs, e.g., there will not be ramp in demand greater than 2 MW/h between 4–8pm

Deterministic forecasts are commonly used today in Transmission operations, with probabilistic and event forecasts growing in adoption in response to the increasing variability and uncertainty from growing shares of renewable penetration. However, as short-term forecasting is comparatively newer to Distribution operations, it is expected that the focus will stay on deterministic forecasts for the near future. Note that some of the lessons learned here could be transferred to support Transmission operations in forecasting of aggregated levels of distributed PV, but the focus of this paper is on Distribution system operations.

Forecast Horizon, Resolution and Schedule

Horizon refers to the time difference between when a forecast is generated and when it applies. For example, a forecast generated by 9am that predicts the load at 10am would be considered an hour-ahead forecast. One common way to categorize forecast horizons is as follows:

- Nowcast: close to real-time (typically less than 5-min ahead)
- Intra-hour: less than 1-hour ahead
- Intra-day: 1-hour to multiple hours ahead
- Day-ahead: 24-hours ahead or longer
- Week-ahead: 7-days ahead

Related to horizon is the time resolution of forecasted values, e.g., hourly vs 15-min averages, and the schedule, i.e., the timeline for how often and at which time forecasts are updated. As with forecast horizon, the forecast resolution and schedule can vary, but they are generally influenced by a combination of physical, computational, and human factors related to the given application. In general, forecasts with shorter horizons that are updated more frequently, e.g., every hour instead of once per day, are more accurate than forecasts with longer horizons or less frequent updates. Similarly, forecasts at coarser time resolutions, e.g., hourly averages instead of 5-min, tend to be more accurate. Therefore, the choice of horizon, resolution and schedule is generally a balance between wanting more accurate forecasts and the inherent constraints of the specified application. Figure 2 illustrates the difference between two schedules: forecasting once per day for the next 24 hours vs forecasting every hour for the next hour.

Note, however, that there can be significant differences in definitions between utilities and applications. For example, day-ahead can be used to refer to forecasts made for a single period 24-hours ahead, the time range covering 24–48 hours ahead or, as in the case of ISO day-ahead markets, forecasts generated during day D that apply to the entire next day D+1 (midnight-to-midnight).

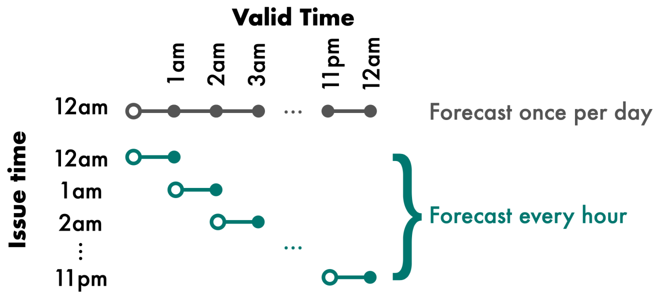

Figure 2. Example of two different forecast schedules.

Forecast Granularity

Granularity refers to the functional representation in the network. Forecasts at finer granularities, e.g., node/customer level rather than at the feeder head, can provide more insight. However, finer granularities can be more challenging to forecast accurately compared to coarser aggregations. Additionally, finer granularity forecasts can require more complex data intake and processing for the forecast user.

Note that granularity is separate but related to spatial resolution, where spatial resolution defines the precision of a model in terms of physical area. For example, consider the problem of forecasting PV production for a feeder with PV installations spread over 10 km^2. In this case, 10 km^2 resolution solar irradiance data may be sufficient for a feeder-level forecast. However, finer spatial resolution data, e.g., 1 km^2, would better capture variations between sites and therefore enable more accurate node-level forecasts.

Forecast Methods and Data Sources

There are many methods for short-term forecasting, each with their own pros and cons. These methods can be generally classified as either a) statistical or b) physics based.

Statistical (or data-driven) methods refer to techniques that learn to predict from data, e.g., Machine Learning models such as Neural Networks, and are among the most common approaches employed today for state-of-the-art forecast performance. Unfortunately, such methods require access to historical measurements of the variable being predicted, e.g., DER production, to train the underlying model and therefore are not viable for scenarios in which measurement data is unavailable.

In contrast, physics-based methods rely on modeling of the underlying physical phenomenon and therefore can be useful in situations where measurements of the target variable are unavailable. For example, the power output of a PV system is primarily a function of solar irradiance. Therefore, given a forecast of solar irradiance at time *t* and knowledge of the PV system, e.g., AC capacity, orientation, and panel tilt, the PV production at time *t* can be estimated using a (deterministic) resource-to-power model. Figure 3 illustrates common sources of solar irradiance data (observations or forecasts) that can be used for predicting PV production:

- **Ground sensors**: instruments installed that directly measure solar irradiance; can provide high quality and high frequency data; not commonly available in all areas.
- **Sky cameras**: wide-angle lens cameras oriented to take pictures of the sky; usually co-located with irradiance sensors; sequences of images can be used to predict cloud motion, which in turn predicts how the solar irradiance will change; best suited for predictions less than 1-hour ahead; less commonly available than ground sensors.
- **Remote sensing**: estimates of solar irradiance from geostationary satellite imagery and physics-based modeling; estimates available for most of the globe, including any point across the United States; current generation of satellites for the US allows irradiance estimates at 5-min, ~1 km^2 resolution.

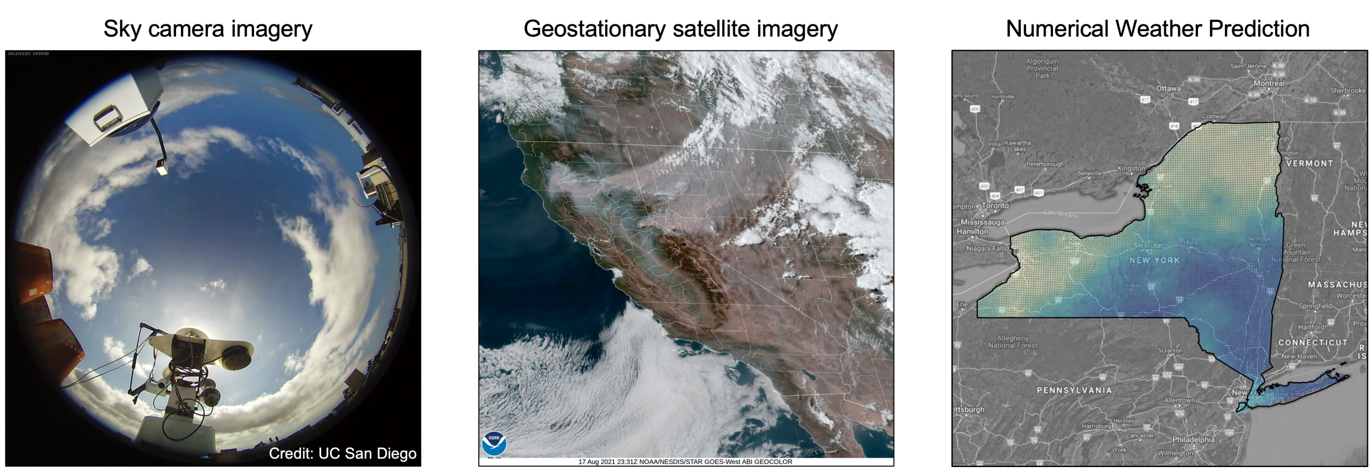

Figure 3. Examples of data sources that can be used in forecasting solar irradiance and therefore PV production.

- **Numerical Weather Prediction (NWP)**: physics-based numerical models that predict a range of meteorological variables including solar irradiance; data can be used as inputs to other models or as forecasts directly; details vary between models, but all run on regular schedules and provide estimates across a gridded domain; generally lower resolution (temporal + spatial) than individual sensors (e.g., hourly and ~12 km^2).

Of the above, remote sensing and NWP sources provide the most extensive coverage and options for the US. Unsurprisingly, many forecast methods and products from commercial vendors use some combination of remote sensing and NWP data, with statistical models used throughout. Additionally, methods based on remote sensing tend to be more accurate for horizons less than ~4–6 hours ahead, with NWP models being used for horizons of ~6-hours to ~7-days ahead. Beyond that horizon, climatological forecasts, perhaps adjusted for seasonal trends, typically perform best.

does not rely on any PV production measurements and therefore can be used to predict for any site where the PV system details are known, e.g., installed capacity, DC:AC ratio, orientation, panel tilt and site shading. Figure 4 illustrates the forecast methodology.

Figure 5 illustrates the diversity of the PV systems considered in the study, as provided by the Utility. The systems vary in size from ~2–12 kW, with a median size of 5.5 kW. While the 'optimal' orientation to maximize energy production is due south for PV in the Northern Hemisphere, constraints on roof mounting can result in a wide range of orientations. Similarly, roof pitch directly influences panel tilt, leading to tilt angles of ~5 degrees to 40 degrees from horizontal.

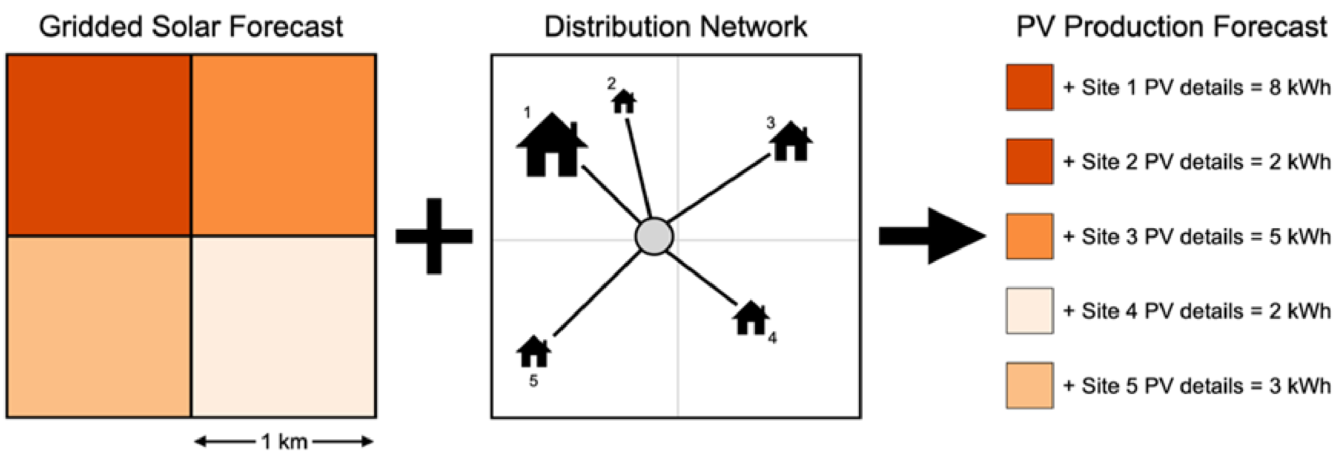

Figure 4. Illustration of using gridded solar irradiance forecasts to predict site-level PV production.
Note: The five sites are connected to the same feeder, but have varying PV system details (AC capacity, panel orientation, etc.) and are physically located within different cells of the 1 km^2 resolution solar forecast grid.

Utility Field Experience

This section summarizes learnings from on-going work between EPRI, a large US utility (hereafter referred to as Utility), and a commercial solar forecast provider (hereafter referred to as Solar Forecaster).

Datasets

The Utility provided approximately two years (January 2019–October 2020) of 15-min resolution direct measurements of PV production from DER-specific meters installed at over 70 sites within the Utility's service territory. The Solar Forecaster provided 15-min resolution PV production forecasts for each site for the same time range as the measurements, with forecast horizons ranging from nowcasts to 7-days ahead. It is important to note that the forecast methodology employed by the Solar Forecaster

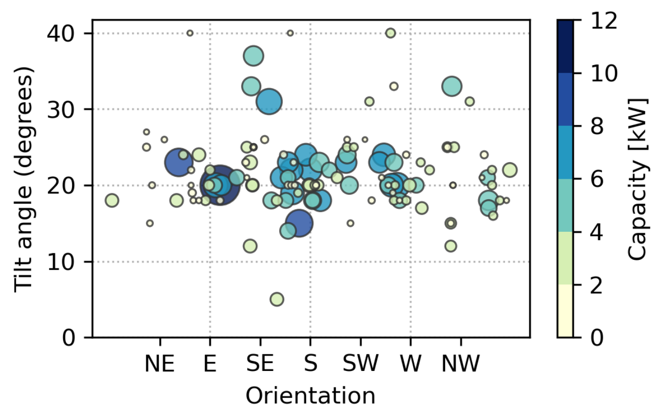

Figure 5. Installed capacity, orientation, and panel tilt (in degrees from horizontal) for the PV systems considered in the study.

Measuring Forecast Performance

There are a multitude of approaches for quantifying forecast performance. Here the focus is on two standard metrics: root mean square error (RMSE) and mean bias error (MBE). RMSE measures the absolute error, with RMSE = 0 for a perfect forecast, while MBE indicates whether a forecast tends to over-predict (MBE > 0) or under-predict (MBE < 0). To enable comparisons between systems of different sizes, all RMSE and MBE values are normalized by the peak PV production for that system based on the installed capacity. Unless otherwise noted, all results are based on the approximately two years (2019–2020) of overlapping forecasts and actuals. For additional information on forecast metrics and evaluations, readers are referred to the Solar Forecast Arbiter project, which provides an open-source framework for forecast evaluations that are "impartial, repeatable, and auditable."[1]

Overall Forecast Performance

Figure 6 illustrates sample forecasts for a single site vs aggregated at the feeder-level. The shorter horizon and more frequent updates enable the nowcasts to better estimate PV production than the day-ahead forecasts, although the day-ahead forecasts still capture the overall behavior. Across all sites, the nowcasts have an average per-site RMSE of 7.2% vs 10.1% for the day-ahead forecasts. Meanwhile, both the nowcasts and forecasts exhibit low bias, with average per-site MBE values of -0.5% and 0.2%, respectively.

Whether these error magnitudes are sufficiently small for Distribution operation applications such as VVO and automated restoration is an open research question. However, the values are in line with published solar forecasting literature and indicate that it is possible to forecast PV production for sites without measured data. For example, Zhang et al. (2014) reported RMSE of 5–12% for intra-hour forecasts of PV production.[2]

Figure 7 shows how error varies between site-level forecasts. More specifically, most sites have RMSE less than 10%, with some sites tending to over-predict (MBE > 0) and others tending to under-predict (MBE < 0). However, there are some outliers with much larger error values due to inaccuracies in the recorded PV system details (sites A and B) and seasonal shading (site C) provided to the forecast model. These outliers emphasize the importance of accurate site information for site-level forecasting when measurement data is unavailable.

Impact of Forecast Granularity

Figure 7 also shows how forecasting aggregations of PV systems (feeder-level) can result in lower error than forecasting individual sites; the feeder-level forecasts achieve RMSE of ~4% and MBE of ~0%. There are two main causes for this trend. First, aggregating the spatially distributed PV systems reduces the variability of total PV production and lower variability profiles are easier to predict. Second, site-level forecast errors may cancel each other out. For example, mis-predicting the timing of a passing cloud may result in over-predicting the solar irradiance–and therefore the PV production–at one site while under-predicting at a neighboring site at the same time, resulting in a net error of zero between the two sites.

Impact of Forecast Schedule

Issuing forecasts once per day for the next 24-hours (daily issued day-ahead) results in a site-level RMSE of 8.9% ± 0.01%, while issuing forecasts every hour for only the next hour (hourly issued hour-ahead)

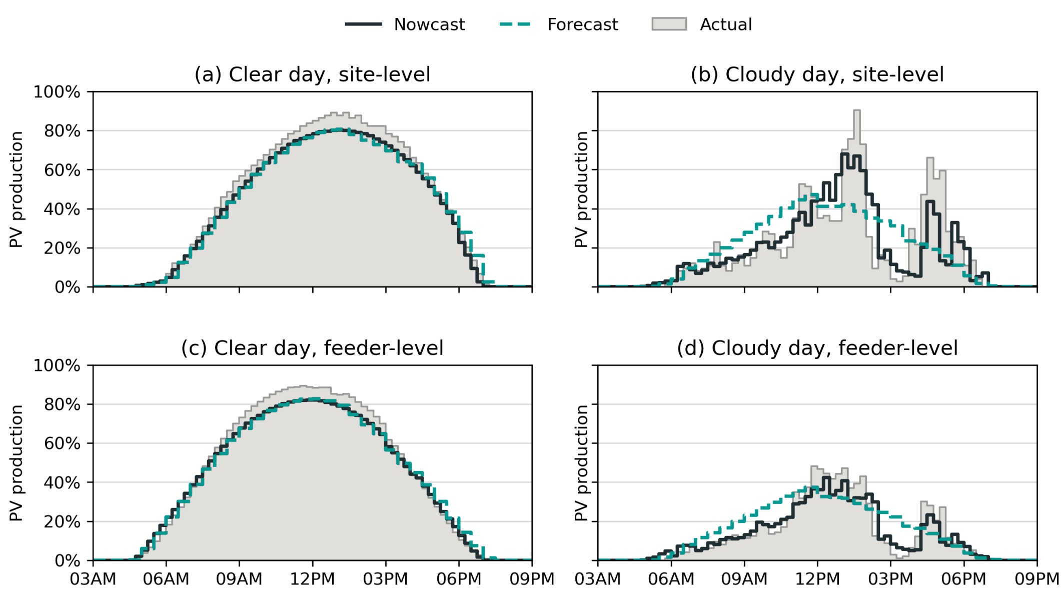

Figure 6. Comparison of nowcasts and day-ahead forecasts against actual PV production for a single site vs at the feeder-level.

[1] Solar Forecast Arbiter: https://solarforecastarbiter.org/.

[2] Zhang et al. (2014) "Baseline and Target Values for Regional and Point PV Power Forecasts: Toward Improved Solar Forecasting", Solar Energy (122), pp. 804–819.

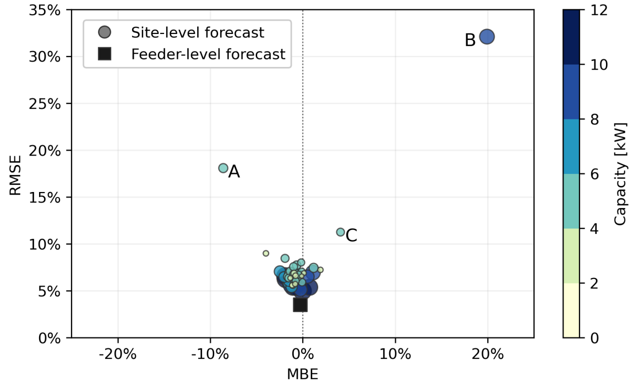

Figure 7. Distribution of RMSE vs MBE for site-level and feeder-level forecasts. Three outlier sites are labelled (A, B and C).

achieves RMSE values of 8.5% ± 0.01%, all else being equal. Following the same trend, feeder-level forecasts result in RMSE values of 6.5% vs 6.1% for the daily-issued and hourly-issued schedules, respectively. Note that the RMSE values reported here are different than in the prior sections due to differences in the merged forecast datasets analyzed and therefore readers should only compare values within the same section, e.g., 8.9% vs 8.5% for daily vs hourly issued site-level forecasts.

As expected, updating forecasts more frequently and with shorter horizons (hourly schedule) enables more accurate forecasts. At the same time, the higher error in the daily schedule forecasts can be offset by practical benefits, e.g., less complex decision making (once per day vs 24 times per day) and more lead time before events of concern, such as large ramps from a passing cloud field. With that said, Distribution operators could also adopt multiple-cycle schedules, e.g., to include both daily and hourly issued forecasts, which is a common practice in Transmission operations.

Impact of Time of Day

Figure 8 shows how forecast error varies by hour of the day. The RMSE increases during the middle of the day when PV production is greatest, with the forecasts tending to under-predict (MBE < 0) during that time. This trend in MBE may be the result of inaccuracies in the PV system details provided to the forecast model. For example, if a 6 kW site was recorded as 5 kW, then the forecasts would tend to under-predict the PV production by 1 kW or more.

Impact of Forecast Horizon

Figure 9 shows how forecast error increases with horizon, with average per-site RMSE growing from ~5% for nowcasts to ~12% for 6-days ahead. The same trend holds for feeder-level forecasts. While not unexpected, these results will be important when considering forecasts for different Distribution operations applications. For example, volt-var optimization VVO using intra-hour or intra-day forecasts should lead to better results than day-ahead forecasts due to the lower error.

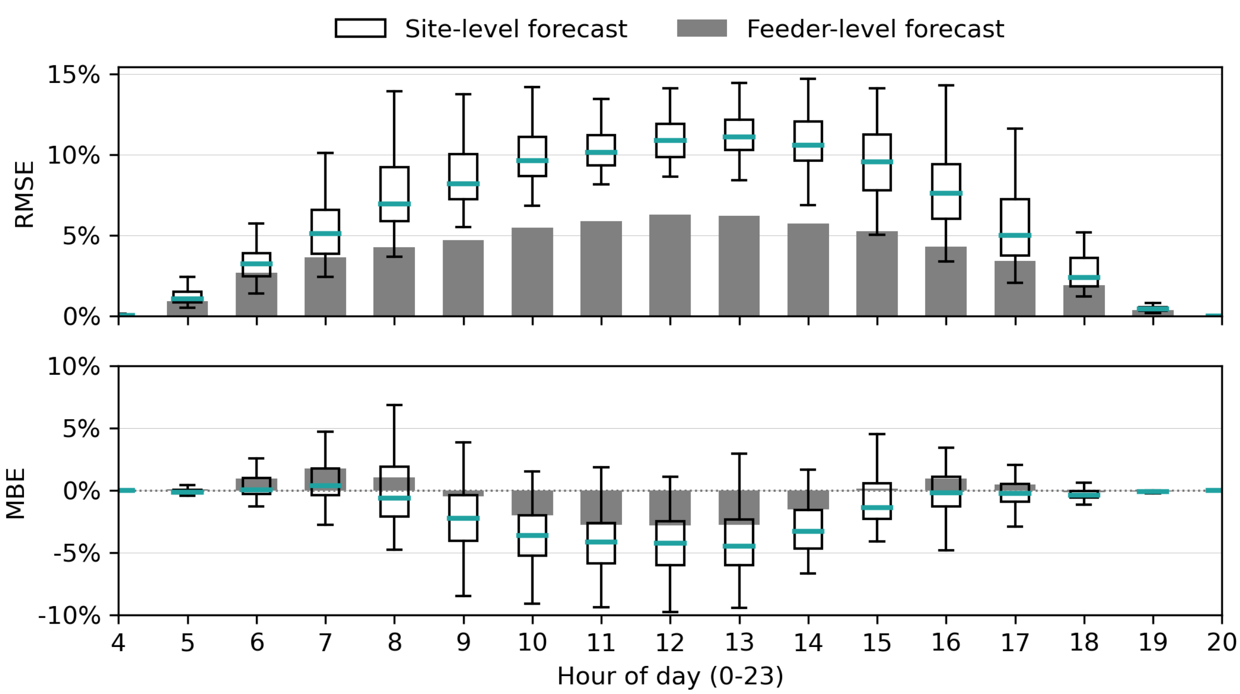

Figure 8. Forecast error by hour of day for site-level and feeder-level forecasts of PV production.

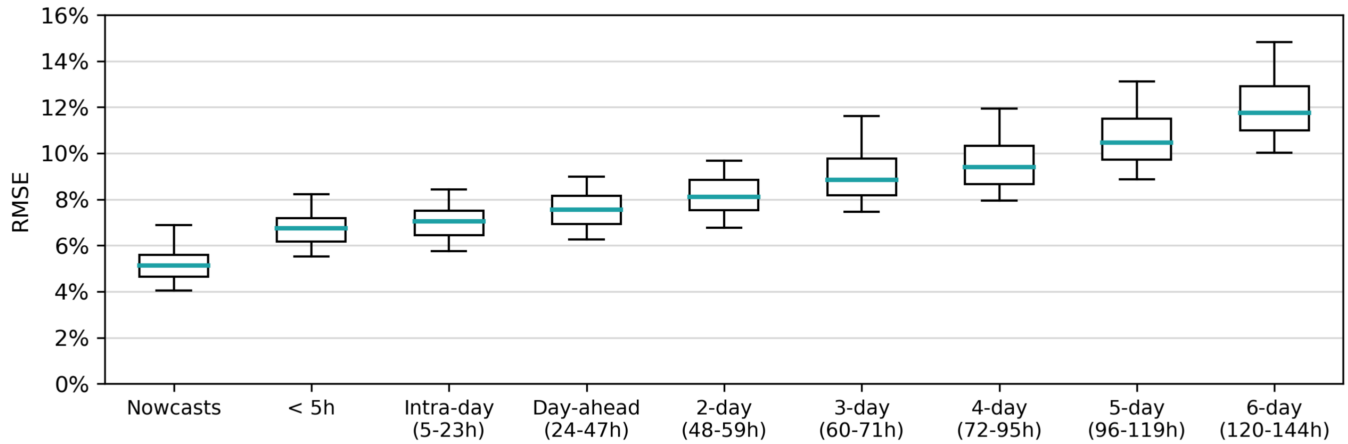

Figure 9. Forecast error as a function of horizon for site-level forecasts.

Improving Forecasts via Measurement Data

While the previous sections have demonstrated the ability of modern forecasting methods to predict PV production without having access to measurement data, a natural question is whether measurement data can further improve forecast performance. More specifically, can feeder-level forecasts be improved if measurement data is available for a subset of sites?

Figure 10 shows how feeder-level forecast error decreases as a function of the number of sites with (historical) measurement data available used to post-process the forecasts, with improvements possible even with only 10% of sites providing measurements. Though the improvements start to level off beyond ~50% of sites with measurement data.

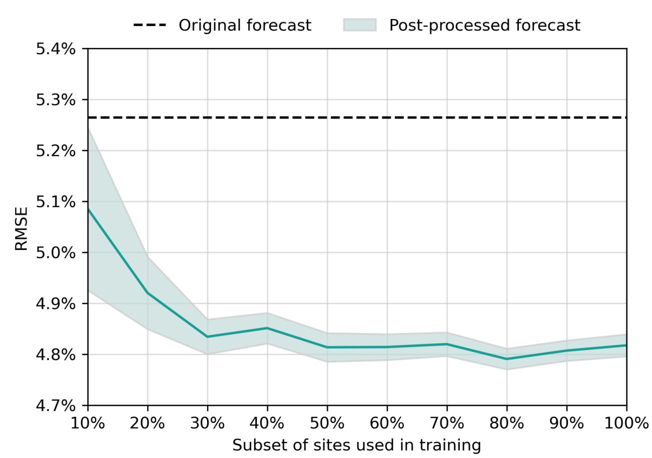

Figure 10. Feeder-level forecast error as a function of percent of sites with measurement data that can be used to post-process the feeder-level forecasts, with the area representing a 90% confidence interval (CI) based on 60 bootstrap simulations.

Conclusions

DER visibility is a growing challenge for Distribution operators as the penetration of PV and other DERs continue to increase. Short-term forecasting offers a potential solution that is both cost-effective and scalable. Results from on-going work with a large US utility and commercial forecast vendor support the following conclusions:

- Modern short-term forecasting methods can provide DER visibility with some degree of accuracy across all timescales of interest to Distribution operations (minutes to hours to days ahead) and without requiring access to any DER measurement data. However, there are open questions regarding the required level of forecast accuracy for different applications, e.g., DER dispatch, hosting capacity and volt-var optimization (VVO).
- Forecast performance depends on the accuracy of DER system details provided to the forecast model, including environmental conditions than impact production, e.g., shading of PV panels from trees and other foliage.
- Forecasting at coarser granularities can be more accurate, e.g., aggregating site-level forecasts to produce a feeder-level forecast can exhibit lower relative error than the individual site-level forecasts. However, the difference in forecast performance does not preclude the use of site-level forecasts in Distribution operations.
- Forecast performance can be improved if (historical) measurement data is available at even 10% of the DER sites being forecasted, though results improve with more sites and the error reduction depends on which sites are chosen and the post-processing methods employed. Additionally, improvements start to level off when measurements are available for greater than ~50% of DER sites.

Based on the above, the next logical step is integrating DER forecasting into short-term load forecasting, as well as using DER forecasts for load disaggregation. A forthcoming technical brief will address both load forecasting and load disaggregation, and will build upon the DER forecasting work presented in this document.

Contact Information

For more information contact the EPRI Customer Assistance Center at 800.313.3774 (askepri@epri.com).

EPRI Resources

Mobolaji Bello, *Technical Leader*
865.218.8005, mbello@epri.com

David Larson, *Engineer/Scientist III*
650.855.2318, dlarson@epri.com

The authors would like to thank Lysha Matsunobu, Aidan Tuohy, Lindsey Rogers, and Brian Deaver for their contributions to this document.

Distribution Operations and Planning

NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or e-mail askepri@epri.com.

The Electric Power Research Institute, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI also provides technology, policy and economic analyses to drive long-range research and development planning, and supports research in emerging technologies. EPRI members represent 90% of the electricity generated and delivered in the United States with international participation extending to nearly 40 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; Dallas, Texas; Lenox, Mass.; and Washington, D.C.

Together . . . Shaping the Future of Energy™

3002021931 October 2021

EPRI
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com